AD-A031 399 TEXAS UNIV AT AUSTIN CENTER FOR CYBERNETIC STUDIES F/G 12/2

LEAST ABSOLUTE VALUE ESTIMATORS FOR ONE-WAY AND TWO-WAY TABLES.(U)

JUN 76 R D ARMSTRONG, E L FROME N00014-75-C-0569

UNCLASSIFIED CCS-266 NL

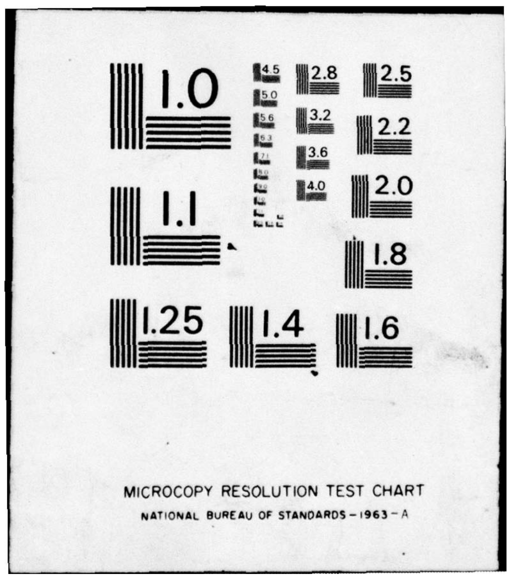
|OF|

AD A031399

END

DATE FILMED

12-76


1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

12

B.S.

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712

Research Report CCS 266

LEAST ABSOLUTE VALUE ESTIMATORS FOR ONE-WAY and TWO-WAY TABLES

by

R. D. Armstrong
E. L. Frome

June 1976

This research was partly supported by Project NR047-021, ONR Contract N00014-75-C-0569 with the Center for Cybernetic Studies, The University of Texas. Reproduction in whole or in part is permitted for any purpose of the United States Government.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 203E
The University of Texas
Austin, Texas 78712
(512) 471-1821

LEAST ABSOLUTE VALUE ESTIMATORS FOR ONE-WAY and TWO-WAY TABLES

by

R. D. Armstrong
E. L. Frome

This paper concerns itself with the problem of estimating the parameters in a one-way and two-way classification model by minimizing the sum of the absolute deviations of the regression function from the observed points. The one-way model reduces to obtaining a set of medians from which optimal parameters can be obtained by simple arithmetic manipulations. The two-way model is transformed into a specially structured linear programming problem and two algorithms are presented to solve this problem. The occurrence of alternative optimal solutions in both models is discussed, and numerical examples are presented.

1. INTRODUCTION

An important problem in statistics is to study the effect of one or two factors on a dependent variable. This problem can be formulated as a regression analysis using dummy (0,1) variables to represent the levels of the factors, and the resulting least squares analysis (LSQ) is well known [29]. Recently, the least squares approach has come under considerable criticism, and several resistant estimation procedures have been proposed [1, 19, 20, 21,22]. Along with the resistant estimation techniques has come an increased computational burden, and in some cases subjective decisions concerning outliers [14,27,31] weight functions [6], and score functions [21] must be made by the statistician. Minimizing the sum of the absolute value of the residuals is a robust procedure [19] which in some cases bypasses the latter difficulty. The computations involved in obtaining least absolute value (LAV) estimates have been a major deterrent to its use. This paper will demonstrate how LAV estimates can efficiently be obtained for one-way and two-way tables. A second difficulty in LAV estimation is the existence of alternate optimal solutions. For one-way and two-way tables alternative optimal solutions will frequently exist. However, a unique solution can be obtained by placing appropriate restrictions on the parameters. From a data analytic point of view this may be regarded as an advantage since it requires some careful thought in selecting additional contraints that will yield a "good" solution. In some situations this simple row-plus-column fit provides a first step in data analysis. The fitted values and residuals are used to identify possible outliers or to suggest how an improved fit may be obtained (see e.g. [1], [6]). The LAV estimates will also provide a good starting point for resistant procedures that are iterative and require residuals from an initial fit to initiate the procedure.

Charnes, Cooper and Ferguson [12] appear to be the first to present a practical approach to solve the general linear LAV problem. They demonstrated how the problem could be transformed into a linear programming problem and thus solved by using the well developed theory of linear programming (LP). They also proved the statistical consistency of the estimates for LAV or any other norm-functional. Other papers primarily concerned with the use of LP to solve LAV problems are [2,4,5,28,31,34]. The main point to be gleaned from the more recent of these references is that a special purpose primal simplex algorithm has proven to be the most efficient method for solving linear LAV problems. A reasonable alternative to the special purpose primal algorithm is to take the dual of the original LP problem and solve it via an LP code with simple upper bounding. Section 3 outlines this transformation for a two factor model while Wagner [32] gives a more detailed presentation for the general case. Computational results [3] indicate that the dual approach takes approximately four times as long as the special purpose primal algorithm, but the algorithm for solving the dual has the definite advantage of being more widely available.

In Section 2, we demonstrate how LAV estimates can be obtained for a one factor model without LP. In Section 3, two computer-oriented approaches for the analysis of a two factor model using LP are presented. Both methods exploit the topological structure of LP problem to provide efficient solution techniques. Section 4 presents sufficient conditions for alternative optimal solutions to exist when additional criteria are not present. Examples illustrating LAV estimation for one-way and two-way tables are given in Section 5.

2. ONE-WAY CLASSIFICATION MODEL

Suppose it is hypothesized that observed values of a random variable are affected by t levels of a certain factor. A statistical model to study these effects may be stated as follows:

$y_{ik} = \mu + \tau_i + e_{ik}$ $i=1, 2, \ldots, t$; $k = 1, 2, \ldots, n_i$, where y_{ik} is the k-th observation at the i-th level, μ is a typical value, τ_i is the effect associated with the i-th level and e_{ik} is an unobservable random "error."

The LAV estimates for μ and τ_i, $i = 1, 2, \ldots, t$ by definition solve the following problem:

$$\text{Minimize } z = \sum_{i=1}^{t} \sum_{k=1}^{n_i} |y_{ik} - (\mu + \tau_i)| \tag{2.1}$$

An immediate difficulty arises because we have one degree of freedom in choosing values for the parameters; that is, μ or any one of the τ_i's may be assigned an arbitrary value without affecting the optimal value of z. The same difficulty arises in LSQ estimation, and is averted by assuming that the total of the effects should be zero. Thus, the degree of freedom is absorbed by the constraint:

$$\sum_{i=1}^{t} \tau_i = 0 \tag{2.2}$$

In a LAV analysis, this degree of freedom must also be removed by an additional constraint, but now the form of the constraint is not so obvious. To see this, consider the t disjoint problems:

$$\text{Minimize } \sum_{k=1}^{n_i} |y_{ik} - \alpha_i|, \; i = 1, 2, \ldots, t, \tag{2.3}$$

where $\alpha_i = \mu + \tau_i$, $i = 1, \ldots, t$.

An optimal value for α_i, say $\tilde{\alpha}_i$, is the median of the points y_{ik}, k=1, 2, ..., n_i. It then follows that if we were using (2.2) as a constraint on the τ_i's, the optimal solution would be:

$$\tilde{\mu} = \left(\sum_{i=1}^{t} \tilde{\alpha}_i \right)/t \tag{2.4}$$

$$\text{and} \quad \tilde{\tau}_i = \tilde{\alpha}_i - \tilde{\mu},\ i = 1,2,\ldots t. \tag{2.5}$$

The $\tilde{\mu}$ given by (2.4) is the arithmetic mean of t medians. One reasonable alternative would be to choose $\tilde{\mu}$ to be the median of all observations, but to parallel the LSQ analysis as closely as possible, we take a different approach. First note that (2.2) is equivalent to taking up the degree of freedom by choosing μ and τ_i, i=1, 2, ..., t so as to minimize

$$\sum_{i=1}^{t} \tau_i^2 = \sum_{i=1}^{t} (\alpha_i - \mu)^2 ,$$

while still providing LSQ estimates. Correspondingly, to obtain parameters for the LAV estimate we minimize

$$\sum_{i=1}^{t} |\tau_i| = \sum_{i=1}^{t} |\alpha_i - \mu| \tag{2.6}$$

while maintaining the minimum value for z.

When the optimal value of α_i is unique (i.e., the median of the points y_{ik}, k=1, 2, ..., n_i is unique) for all i, $\tilde{\mu}$ is the median of $\tilde{\alpha}_i$, i=1, 2, ..., t and $\tilde{\tau}_i$ is obtained from (2.5). However, frequently the median of the y_{ik}'s is not unique but rather can lie anywhere within a continuous closed interval. When this is the case, $\tilde{\mu}$ can be obtained as follows.

<u>Step 1.</u> Set U equal to the smallest lower bound of the intervals within which the optimal value of the α_i's must lie (unique values have the same upper and lower bound.)

<u>Step 2.</u> Increase the value of U until any further increase would place more intervals completely below U than there are completely above U.

<u>Step 3.</u> Place L equal to the current value of U.

<u>Step 4</u>. Decrease the value of L until any further decrease would place more intervals completely above L than there are completely below L.

All the values in the closed interval [L, U] are optimal for μ subject to the additional criterion (2.6). Let μ^* denote the median of all the y_{ik}'s and choose the point in [L, U] that minimizes $|\mu-\mu^*|$. This criteria provides an estimate that is as close as possible to the estimate of μ that is obtained under the minimal one parameter model (i.e. under they hypothesis that all the τ_i's are zero). A similar procedure will be used in obtaining estimates of the parameters in the two factor model (see section 5). Once $\tilde{\mu}$ has been chosen from within this interval, $\tilde{\alpha}_i$ is chosen to be as close to $\tilde{\mu}$ as possible while remaining in the range of optimality for (2.3). The τ_i's are then determined from (2.5).

This LAV estimate is unique. Although the additional criteria that were added to force a unique solution are arbitrary, they are reasonable for this situation. Other approaches, similar to this goal programming (constrained regression) approach [8, 9, 11, 24] proposed here, can be used to define a unique optimal solution or we could let $\tilde{\mu} = (L + U)/2$. Unless these additional criteria are rather complex, LAV estimates are easily obtained for a one-way table. However, as we shall see in the next section, the extension of the LAV approach to two-way tables is far more complex than the corresponding LSQ extension.

3. TWO-WAY TABLE

3.1 Definition of Model

A two factor model arises when a second factor is introduced as follows:

$$y_{ijk} = \mu + \tau_i + \beta_j + e_{ijk}, \quad \begin{array}{l} i=1, \ldots, r; \\ j=1, \ldots, c; \\ k=1, \ldots, n_{ij}. \end{array}$$

Thus, y_{ijk} is the k-th observation at the i-th level of the first factor and the j-th level of the second factor; τ_i represents the effect of the i-th level of the first factor (i.e. row effect), β_j represents the effect of the j-th level of the second factor (column effect), and μ is a typical value. LAV estimates of the parameters are obtained by solving the following problem:

$$\text{Minimize } z = \Sigma_{i=1}^{r} \Sigma_{j=1}^{c} \Sigma_{k=1}^{n_{ij}} | \; y_{ijk} - (\mu + \tau_i + \beta_j) \; |. \tag{3.1}$$

There are two degrees of freedom in assigning values to μ, τ_i, and β_j; thus, restrictions should be added to the problem. In the corresponding LSQ analysis (2.2) and

$$\Sigma_{j=1}^{c} \beta_j = 0 \tag{3.2}$$

are appended. This is equivalent to providing LSQ estimates which minimize

$$\Sigma_{i=1}^{r} \tau_i^2 + \Sigma_{j=1}^{c} \beta_j^2$$

Analogously, in the LAV analysis we provide estimates that minimize

$$\Sigma_{i=1}^{r} |\tau_i| + \Sigma_{j=1}^{c} |\beta_j| \tag{3.3}$$

subject to the optimal value for z in (3.1) being maintained. As in the one-way analysis, this additional criterion does not necessarily provide a unique solution (see Section 5 for an alternative). Further restrictions, or a completely different set of criteria, may determine a unique solution.

Our purpose is to present what we feel are reasonable conditions for LAD estimates to satisfy.

3.2 Computational Approaches

Before (3.3) is considered, two computer oriented approaches to obtain an optimal solution to (3.1) will be discussed. We again make the transformation $\alpha_i = \mu + \tau_i$, and restate (3.1) as:

$$\text{Minimize } z = \Sigma_i \Sigma_j \Sigma_k \mid y_{ijk} - (\alpha_i + \beta_j) \mid, \tag{3.4}$$

Problem (3.4) can be written as a linear programming problem

$$\text{Minimize } \Sigma_i \Sigma_j \Sigma_k (d^+_{ijk} + d^-_{ijk}), \tag{3.5}$$

subject to:

$$\alpha_i + \beta_j - y_{ijk} - d^+_{ijk} + d^-_{ijk} = 0,$$

$$d^+_{ijk} \geq 0,\ d^-_{ijk} \geq 0,$$

$$i=1, \ldots, r; \quad j=1, \ldots, c; \quad k=1, \ldots, n_{ij},$$

where d^+_{ijk} and d^-_{ijk} are the positive and negative deviation of the regression equation from the observation y_{ijk}, respectively. Problem (3.5) is not tractable in its present form for a direct application of the simplex algorithm. The main reason for this is that the number of constraints is equal to the number of observations which may give rise to an excessively large basis matrix. This difficulty can be overcome by taking the dual of (3.5) which is given by:

$$\text{Maximize } \Sigma_i \Sigma_j \Sigma_k\, y_{ijk}\, \pi_{ijk} \tag{3.6}$$

subject to:

$$\Sigma_j \Sigma_k \pi_{ijk} = 0\ ,\ i=1, \ldots, r;$$

$$\Sigma_i \Sigma_k \pi_{ijk} = 0\ ,\ j=1, \ldots, c;$$

$$-1 \leq \pi_{ijk} \leq 1\ ,\ i=1, \ldots, r; \quad j=1, \ldots, c; \quad k=1, \ldots, n_{ij}.$$

By making the transformation $\pi'_{ijk} = \pi_{ijk} + 1$ (3.6) can be written in a more standard linear programming format

$$\text{Maximize } \Sigma_i \Sigma_j \Sigma_k (y_{ijk}\, \pi'_{ijk} - y_{ijk}), \tag{3.7}$$

subject to:

$$\Sigma_j \Sigma_k \pi'_{ijk} = \Sigma_j n_{ij} \text{ , i=1, ..., r,}$$

$$\Sigma_i \Sigma_k \pi'_{ijk} = \Sigma_i n_{ij} \text{ , j=1, ..., c,}$$

$$\text{and} \quad 0 \leq \pi'_{ijk} \leq 2 \text{ , i=1, ..., r; j=1, ..., c; k=1, ..., } n_{ij}.$$

It can now be recognized that (3.7) is a capacitated transportation problem [10] with r origins and c destinations except that, because of multiple observations within cells, there is more than one path from origin i to destination j. This extension can be incorporated into the standard LP algorithm by only considering π'_{ijk}'s for entry into the basis when all other LP variables corresponding to observations in cell (i,j) with a value larger than y_{ijk} are at their upper bound. Computational results [15,16] indicate that transportation problems can be solved approximately 150 times faster by using a special purpose primal simplex code as opposed to a general purpose state-of-the-art LP code. Thus, considerable savings can be derived by recognizing the special structure of (3.6).

Once (3.6) has been solved, optimal values for the α_is and β_js in (3.4) are given by the dual variables or simplex multipliers for the first r + c constraints. There is, however, one degree of freedom in choosing the α_i s and β_j's, and a second degree of freedom in assigning values to the τ_is and μ. These degrees of freedom can be taken up by satisfying criterion (3.3). We delay the discussion of how to accomodate (3.3) until after the primal approach to (3.5) has been presented.

For the general problem of obtaining parameters for absolute deviations estimates, it has been shown [3, 4, 30] that solving the general case equivalent of (3.5) directly with a special purpose primal algorithm is computationally the most efficient approach available. There is no reason to believe that this would not also be true here as the structure of the problem can still be utilized to perform the operations of the algorithm without ever inverting a matrix explicitly. This algorithm, modified to take advantage of (3.5)'s structure, will not be developed here, but a brief overview is given to indicate the use of techniques found in solving transportation problems and to state a formula required in the next section.

We begin by restating the constraints of (3.5) in matrix notation as follows:

$$A\gamma - Y - D^+ + D^- = 0, \tag{3.8}$$

$$D^+ \geq 0,\ D^- \geq 0, \tag{3.9}$$

where $\gamma' = (\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_c)$, and A is an $(\Sigma_i \Sigma_j n_{ij})$ by (r+c) matrix of 0's and 1's with a single dependent column. It is clear from the objective function of (3.5) that (3.8) can also be written as:

$$Y - D^+ \leq A\gamma \leq Y + D^-. \tag{3.10}$$

The algorithm at any stage works with a basis consisting of (r+c-1) rows of A. To distinguish between the basic and nonbasis rows of A we partition A, Y, D^+, and D^-, and rewrite (3.10) as

$$\begin{bmatrix} Y_b \\ Y_n \end{bmatrix} - \begin{bmatrix} D_b^+ \\ D_n^+ \end{bmatrix} \leq \begin{bmatrix} B \\ N \end{bmatrix} \gamma \leq \begin{bmatrix} Y_b \\ Y_n \end{bmatrix} + \begin{bmatrix} D_b^- \\ D_n^- \end{bmatrix},$$

where B denotes the basis and N the remaining rows of A. If we let $\lambda = B\gamma$, then the constraints (3.8) become

$$Y_b - D_b^+ \le \lambda \le Y_b + D_b^-$$
$$Y_n - D_n^+ \le NB^{\#}\lambda \le Y_n + D_n^- \ ,$$

where $B^{\#}$ is a generalized inverse of B. The current solution is $\lambda^* = Y_b$, $D_b^+ = D_b^- = 0$, $\gamma^* = B^{\#}\lambda^*$, and the deviations in the nonbasic rows are as small as possible based on λ^* and the constraints. The structure of B allows it to be stored as a spanning tree [15] similar to that of the basis of a transportation problem. This allows us to use the triangularity of B (after a dependent column is dropped) to solve $B\gamma = \lambda$ without explicitly obtaining a $B^{\#}$. Thus, $B^{\#}$ is not required to obtain a basic solution; and, in fact, is never needed.

The next step in the algorithm is to determine if an increase or decrease in any λ_i away from its current value λ_i^* will decrease the objective value. The objective function rate of change is $1 + \theta_i$ or $1 - \theta_i$ when λ_i is increased or decreased, respectively, where $\theta = (\theta_1, \theta_2, \ldots, \theta_{r+c-1})$ is given by

$$\theta = \Sigma_j^- N_j B^{\#} - \Sigma_j^+ N_j B^{\#} = (\Sigma_j^- N_j - \Sigma_j^+ N_j) B^{\#}$$
$$\text{or} \quad \theta B = \Sigma_j^- N_j - \Sigma_j^+ N_j . \tag{3.11}$$

In (3.11) the + and - superscripts indicate summation over rows of N with positive and negative deviations, respectively, When a nonbasic row has a zero residual it is classified by the algorithm as a positive or negative zero; thus, every nonbasic row will be included exactly once in the above summation. Again the triangularity of B allows us to obtain θ without B# (just as in the transportation algorithm a primal solution is obtained). Equality (3.11) will be important in the next section where conditions for alternative solutions to the LP problem (3.5) are discussed.

Since we are minimizing the sum of the absolute deviations, a basis change would be made if $|\theta_i| > 1$. Hence, the current solution is recognized as optimal when $|\theta_i| \leq 1$, $i = 1, 2, \ldots, r + c - 1$.

The pivot rule of Barrodale and Roberts [4] which may combine several standard simplex pivots into one is used to determine the row of N to enter the basis. The details of implementing this rule while effectively utilizing the structure of B and N will not be given here, but the major computational step is similar to calculating the ratio in a dual simplex transportation algorithm [17].

3.3 A Secondary Criterion for Choosing the Parameters Estimates

The remainder of this section will be devoted to describing how the secondary criterion (3.3) can be considered within the framework of an LP algorithm. The procedure can be utilized on a revised version of (3.5) or (3.7), and is similar to the perturbation method of Charnes [7]. Like Charnes' method an arbitrarily small positive number ε will be used in the description, but the most efficient implementation would never assign a value to ε and all calculations involving ε are performed implicitly. However, by placing ε equal to a specific value additional computer coding is avoided. We begin by noting that (3.3) can be expressed in LP form as

$$\text{Minimize } \Sigma_{j=1}^{r+c} (\overset{+}{\delta}_j + \overset{-}{\delta}_j), \qquad (3.12)$$

subject to:

$$\begin{aligned} &\tau_i - \overset{+}{\delta}_i + \overset{-}{\delta}_i = 0 \ , \ i=1, \ldots, r \\ &\beta_j - \overset{+}{\delta}_{r+j} + \overset{-}{\delta}_{r+j} = 0 \ , \ j=1, \ldots, c \\ &\overset{+}{\delta}_j \geq 0 \text{ and } \overset{-}{\delta}_j \geq 0 \ , \ j=1, \ldots, r+c \ , \end{aligned} \qquad (3.13)$$

where $\overset{+}{\delta}_j$ and $\overset{-}{\delta}_j$ are the positive and negative deviations of the effects from zero.

Problem (3.12) is only of secondary concern, as we wish to always obtain an optimal solution to (3.5). The desired optimal solution is given in a limiting sense ($\varepsilon \to 0$) by solving

$$\text{Minimize } \Sigma_i \Sigma_j \Sigma_k (\overset{+}{d}_{ijk} + \overset{-}{d}_{ijk}) + \Sigma_{j=1}^{r+c} (\varepsilon\overset{+}{\delta}_j + \varepsilon\overset{-}{\delta}_j) \ , \qquad (3.14)$$

subject to the constraints (3.8), (3.9), and (3.13).

Problem (3.14) is not in the form where the columns have the exact structure that the rows of a transportation problem possess. To obtain the desired format let $\beta_{c+1} = -\mu$ and create a "dummy parameter" β_{c+2} (this is a variable in the LP problem). The problem then becomes

$$\text{Minimize } \Sigma_i \Sigma_j \Sigma_k (\overset{+}{d}_{ijk} + \overset{-}{d}_{ijk}) + \Sigma_{j=1}^{r+c} (\varepsilon \overset{+}{\delta}_j + \varepsilon \overset{-}{\delta}_j), \tag{3.15}$$

subject to:

$$\alpha_i + \beta_j - y_{ijk} - \overset{+}{d}_{ijk} + \overset{-}{d}_{ijk} = 0, \quad i=1,\ldots, r, \; j=1,\ldots, c, \; k=1,\ldots, n_{ij},$$

$$\alpha_i + \beta_{c+1} - \overset{+}{\delta}_i + \overset{-}{\delta}_i = 0 \quad , i=1,\ldots, r,$$

$$\beta_{c+2} + \beta_j - \overset{+}{\delta}_{r+j} + \overset{-}{\delta}_{r+j} = 0 \quad , j=1, \ldots, c; k=1, \ldots, n_{ij};$$

$$\overset{+}{\delta}_j \geq 0, \overset{-}{\delta}_j \geq 0, j=1, \ldots, r+c.$$

where the degree of freedom in assigning values to the parameters is absorbed by always placing $\beta_{c+2} = 0$.

The algorithm described to solve (3.5) directly can be used to solve (3.15) with a slight modification to account for a weight of ε, rather than one, on the deviations of τ_i and β_j away from zero. Also, by taking the dual of (3.15) and by making the lower bound on the variables in this dual problem zero, a capacitated transportation is again created. This can, of course, be solved with a standard code; however, care must be taken to ensure $\beta_{c+2} = 0$ when working back to the optimal solution to (3.15).

The formulation just described takes care of the two degrees of freedom at the expense of creating an additional "source" and an additional "destination" in the transportation problem, and the possibility of alternative optimal solutions has been reduced considerably. The problem of alternative optimal solutions to (3.3) is discussed in the next section along with statements of sufficient conditions for alternative optimal solutions to exist.

4. Alternative Optimal Solutions

A disturbing aspect of LA^V estimation for two-way tables is that alternative optimal solutions frequently occur and, decidedly different estimates are obtainable. This difficulty may be averted by specifying additional criteria for the estimates to satisfy. It is the purpose of this section to indicate that alternative optimal "fits" are to be expected in analyzing two-way tables via LA^V procedures if (3.1) is the sole criterion. This serves to emphasize the importance of "good" additional criteria.

It is well known that the median of an even number of observations is unique only when the two middle observations have the same value. The parameters for the one-way model (2.3) are obtained by taking the median of t sets of observations and the values will be unique only when all t medians are unique. However, a unique solution can always be obtained by adding the additional restrictions described in section 2.

With respect to the two-way model, it can be shown that an LP solution to (3.5) is optimal if the θ obtained by solving (3.11) satisfies

$$-1 \leq \theta_i \leq 1, \; i=1, 2, \ldots, r+c-1.$$

Furthermore, the basis B is a unique optimal basis only when

$$-1 < \theta_i < 1, \; i=1, 2, \ldots, r+c-1;$$

in other words, an alternative optimal basis exists if at the completion of the algorithm θ_i equals -1 or +1 for any i. But because N_j is a vector of 0's and 1's, and because B is a unimodular matrix [18], θ will always have integer components. Hence, at optimality, θ_i will equal either -1, +1 or 0. This means that the current optimal basis is unique if and only if all the components of θ are zero, and this will only be true when

$$\overset{-}{\Sigma}_j N_j - \overset{+}{\Sigma}_j N_j = 0 \tag{4.1}$$

Condition (4.1) forms the foundation for proving the theorem of this section. It might be well to point out at this time that all our results relate only to alternative optimal basic matrices--not to alternative optimal fits. However, an alternative optimal basis is equivalent to an alternative optimal fit whenever an optimal fit interpolates exactly $r+c-1$ points. Theoretically, for fixed sample size this will occur with probability one whenever the observations are taken from a continuous population.

The following theorem is concerned with the special case of two-way classification model where $n_{ij}=1$ for all i and j.

Theorem 4.1. The LP problem (3.3), which is equivalent to the problem of finding LAV estimates for a two-way classification model with exactly one observation per cell, will have alternative optimal basic matrices whenever the minimum of r and c is even.

Proof. Suppose that the LP problem has been solved and an optimal basic matrix obtained. For this matrix to be a unique optimal basic matrix, condition (4.1) must be satisfied. This occurs only if for each nonzero component from an N_j associated with a positive deviation there corresponds a nonzero component from an N_j associated with a negative deviation. In other words, the nonbasic rows of A must contain an even number of nonzero coefficients in each column because summing an odd number of plus or minus ones will never yield zero. The proof of the theorem will consist of showing that whenever the minimum of c or r is even, there is at least one column with an odd number of nonzero components (+1's) in the nonbasic rows.

For explanatory purposes, we assume $r \geq c$, but the proof follows in an analogous manner if the reverse is true. There are $r+c-1$ rows of A in the basis B and, because B forms a basis, every column of B must have at least one nonzero entry. It is noted that B is a submatrix of A and each row has one and only one nonzero entry in the first r components, and one and only one non-zero entry in the last c components. Also, each of the last c columns of A has exactly r 1's with the remaining coefficients being 0's. In order to satisfy (4.1) there must be an even number of 1's in each of the last c columns of N. Thus, since r is even, there must be an even number of 1's in these c columns of B. But each column of B must have at least one nonzero entry, there must be at least two 1's present. This would require B to have at least $2c > c + r - 1$ rows. Therefore, at least one of the last c columns of B has a single nonzero entry. The proof of the theorem now follows from the inability to satisfy (4.1).

It is not difficult to derive examples of two-way tables with a unique optimal basis for the LP equivalent. Consider the two-way table of exhibit 1. This example has a unique optimal basis matrix with the optimal fit interpolating observations 3, 4, 5, 6, and 7. However, alternative optimal basic matrices exist if the 6 and 8 interchange position. Certainly our computational experience would indicate that, even if the conditions of theorem 4.1 are not satisfied, alternative optimal fits are more likely to appear than not.

Furthermore, a unique optimal basis matrix does not clearly define the estimates for the parameters. There are two degrees of freedom that provide us with the ability to arbitrarily choose values for two parameters and remain optimal. In the previous sections we have proposed additional criteria that deal with this problem, and we will discuss this matter further in the next section via numerical examples.

5.1 LAV Analysis of One-Way Table

To illustrate the application of the algorithm described in Section 2, we will use the Nebraska voting data shown in Exhibit 2 [31, chap. 19]. In this section two separate one-way analyses will be carried out. These results are used in Section 5.3 where the same data is used to illustrate the LAV analysis of a two-way table.

First, we consider the rows (i.e. counties). The median intervals are shown in Exhibit 3, and using the algorithm in Section 2, we obtain (L,U)= (325,342). Since the median $\mu^*=338$ we set $\mu=338$ and obtain the fitted values and effects shown in the last two columns of Exhibit 3. The same procedure is then applied to the columns (i.e. years) and the fitted values and effects are shown in Exhibit 4.

In the LSQ analysis the best estimates of the row and column effects (along with the overall mean) provided the solution to the two-way analysis. For the LAV analysis this is not true, but in Section 5.3 we propose obtaining the LAV estimates for the two-way table that are as close as possible to these restricted fits. It will then be possible to assess the relative importance of the row effects after column effects have already been included in the model.

5.2. Some Small Examples for Two-Way Tables

To illustrate the difficulties arising in choosing values for the parameters in the two-way model, we consider the table given by exhibit 5. We begin the analysis by obtaining LAV estimates with the additional restrictions $\tau_1 + \tau_2 = 0$ and $\beta_1 + \beta_2 = 0$. Optimal parameter values can be obtained from the LP by considering any extreme point defining a hyper-plane passing through three of the four observations. Thus, four optimal extreme point solutions are possible and are given by exhibit 6. All optimal solutions are given by convex combinations of these four points. Clearly, a great deal of discrepancy is possible among optimal solutions. Considering only extreme point solutions, the observation which the hyperplane does not interpolate will have a residual with absolute value 998 and the other points will, of course, have a zero residual. Thus, the LP solution could indicate any one of the four observations to have an unduly large residual and make it a candidate for consideration as an outlier.

If we perform the LAV analysis on the same table with additional criterion (3.3) rather than (2.2) and (3.2), a unique solution ($\mu = 1$, $\tau_1 = \tau_2 = \beta_1 = \beta_2 = 0$) is obtained. An inspection of the residuals now indicates that the observation in cell (2,2) might be considered as an outlier.

The two by two table of exhibit 5 is an extreme case and was presented to indicate what could occur in the LAV analysis of two-way tables if caution is not exercised. Generally, such widely divergent solutions will not be available--regardless of the additional criteria employed to absorb the degrees of freedom.

The next example (exhibit 7) is a four by four table from Tukey [31, chp. 22]. With the additional criterion (3.3) appended to the problem, eight optimal extreme point solutions were found. These are given in exhibit 8. The complete set of optimal extreme point solutions can be generated (see [18], p. 166); however, the amount of work required to do so is generally prohibitive. No attempt was made to generate all optimal extreme solutions to the two-way table in exhibit 7.

It is noted that all the solutions of exhibit 8 indicate the row effects being small relative to the column effects. Also, the residual for the outlier of cell (3,2) is 270 for all but the last solution where it is 271.

5.3 LAV Analysis of a Two-Way Table

In Section 5.1 we considered the one-way analysis of the Nebraska voting data. The two-way analysis of this data using LSQ is shown in Exhibit 9 . The LAV estimates are obtained by solving (3.1) with the additional criterion that we minimize

$$|\mu-\mu^*| + \sum_{i=1}^{r} |\tau_i - \tau_i^*| + \sum_{j=1}^{c} |\beta_j - \beta_j^*|, \tag{5.1}$$

subject to the optimal value for z in (3.1) being maintained. In (5.1), the * superscript denoted the LAV estimates that are obtained for the one-way fits (see Section 5.). The robust elementary analysis obtained using LAV is shown in Exhibit 10. The stem-and-leaf plots, hinges, midspreads, and side values for the residuals obtained from the LSQ and LAV fits are shown in Exhibit 11 , and the large (outside, i.e. past the side values) residuals are identified in Exhibit 12 .

Tukey [31, chap.19] obtained a resistant elementary analysis of this data using pomedian polishing on the LSQ analysis (see Exhibit 13). The pomedian procedure leads to residuals that are "nearly balanced" in sign in each row and column. Note that the median of each row and column of the LAV residuals in Exhibit 10 is zero.

5.4 Quality of Fit for Two-Way Table

In a LSQ analysis of a two-way table the importance of the row and column effects is measured in terms of the decrease in the sum of squares that occurs when the row (column) effects are included in the model. For the LAV analysis it is also possible to obtain an indication of the importance of the row and column effects. First, we obtain $Z(\mu^*) = \Sigma_i \Sigma_j \; |y_{ij} - \mu^*|$ = 13661 Next calculate $Z(\mu^*,\beta^*) = 6282$, $Z(\mu^*,\tau^*) = 12531$, and $Z(\tilde{\mu},\tau,\beta)$ = 4240. Then using an approach suggested by McNeil and Tukey [26], we determine that the column fit accounts for $100[\,1 - (6282/13661)^2\,] = 78.9\%$

of the total variation, measured on a size-squared scale in terms of the sum of the absolute deviations. Similarity, 15.9% of the total variation is explained by the column Fit, and the row-plus-column fit accounts for 90.4% of the total variation. Thus, we are able to conclude that the size of the residuals is considerably reduced if both row and column effects are included in the model.

As is the case in an unbalanced LSQ analysis the reduction in the objective function that occurs when additional parameters are added to the model is order dependent. Consequently, this heuristic approach to evaluating the relative importance of a given subset of parameters is similar to the use of r^2 values in the LSQ analysis. This approach is suggested when the analysis is essentially exploratory in nature. We are, on the one hand, prepared to obtain evidence that a simpler model may adaquately fit the data. At the same time we are ready to look at the residuals from a robust row-plus-column fit using diagnostic plots or other techniques that could suggest that outliers are present, or that an improved fit may be had by adding terms to the model or by reexpressing the data.

6. Discussion

It is generally recognized that after obtaining a simple row-plus-column fit for a two-way table the residuals should be carefully analyzed. Gentleman and Wilk [13] have considered the effect of one or two outliers superimposed on a basic additive model with independent normal fluctuations with mean zero and constant variance. Their results indicate that when one outlier is present the judicious use of half-normal plotting provides a complete basis for data-analytic judgements. They further find that direct analysis of residuals (from a LSQ fit) is not reliably indicative of the existence of peculiarities when two outliers are present. Gentleman and Wilk [14] have also considered the problem of multiple outliers, and proposed methods for identification of the "K most likely outlier subset" (where the maximum possible value of K must be known). Their approach considers to what extent a p-parameter model analysis can be statistically improved by selective reduction in the size (n) of the data. Their method results in a LSQ analysis of the "good data".

In many situations the form of the model is only tenative and a diagnostic plot of the residuals is required [31]. The diagnostic plot may suggest that an improved fit can be obtained by either adding to the model or reexpressing the y's. When there are several possible departures from the ideal additive model for a two-way table, the importance of obtaining a robust fit is increased if attempts to improve upon the conventional LSQ analysis are to be successful. Thus as McNeil and Tukey [26] have shown, it is possible to begin with a simple row-plus-column fit of a two-way table using both LSQ and LAV. If the unknown e_{ij}'s follow a Gaussian Distribution; then we expect that the residuals from both fits should appear to be near Gaussian, with somewhat less stretched tails for the least square residuals. If the e_{ij}'s are from a tail-stretched distribution, the residuals should be tail-stretched--the LSQ residuals much less than the e_{ij}'s and the LAV residuals slightly more. Tail-stretched residual distributions may also be the result of an inadaquate model. Consequently, if the LSQ and LAV anlayses are clearly different then further careful analysis is required. It is important to note as Mallows [25] has pointed out, that our understanding of robust techniques and the behavior of the residuals that they generate is limited. Certainly, the results presented in this section indicate that good judgement must be applied by the data analyst obtaining a sensible LAV fit.

5	6	7
4	8	1
3	2	9

Exhibit 1. Two-way table with a unique optimal LP basis matrix.

Exhibit 2. Nebraska Voting--Raw % Democratic for

11 Counties in 12 Presidential Elections (unit = .1%)

Year

County	'20		'28		'36		'44		'52		'60	
D0	353	358	589	757	544	365	345	337	189	167	236	396
D1	323	252	236	669	396	267	238	257	149	148	138	290
B1	288	302	305	619	510	397	372	411	234	268	279	389
D2	379	372	270	606	497	363	388	433	196	223	251	399
D4	342	226	264	626	510	407	404	496	230	264	222	374
B4	270	291	247	569	450	354	325	374	218	259	229	410
D5	228	177	150	553	426	349	272	472	177	240	225	336
B5	270	237	227	561	425	352	340	360	179	232	189	310
D6	265	196	165	547	472	336	313	436	195	219	226	388
B7	322	257	454	661	513	384	379	454	253	307	370	462
D7	270	191	352	776	526	463	442	553	337	358	360	439

Source: Tukey, J. W. (1971). Exploratory Data Analysis, II. Addiso-Wesley,

Exhibit 3. LAV One-Way Analysis of Counties

(i.e. rows of Exhibit 2) for the Nebraska Voting Data

County	Median Interval	Fit	Effect
D5	(240,272)	272	-66
D1	(252,257)	257	-81
D6	(265,313)	313	-31
B5	(270,310)	310	-28
B4	(291,325)	325	-13
B1	(305,372)	338	0
D4	(342,374)	374	36
D0	(353,358)	358	20
D7	(360,439)	439	101
D2	(372,379)	379	41
B7	(379,384)	384	46

Note: The Order has been changed to illustrate the Procedure for obtaining the interval (L,U) = (325,342).

Exhibit 4. LAV One-Way Analysis of Years (i.e. columns)

for the Nebraska Voting Data

	Year											
	'20		'28		'36		'44		'52		'60	
effect	-50	-86	-74	281	159	25	7	95	-142	-98	-109	51
fit	288	252	264	619	497	363	345	433	196	240	229	389

1	1
1	999

Exhibit 5. Sample two-way table with a single outlier.

Extreme point solution	μ	τ_1	τ_2	β_1	β_2
1	500	-499	499	0	0
2	1	-499	499	-499	499
3	500	0	0	-499	499
4	1	0	0	0	0

Exhibit 6. Optimal extreme point solutions to the two-way table of exhibit 5 with the constraints $\tau_1 + \tau_2 = 0$ and $\beta_1 + \beta_2 = 0$ added.

718	732	734	793
725	781	725	716
704	1035	763	758
726	765	738	761

Exhibit 7. Sample two-way table from Tukey [28 chp.22].

Solution	μ	τ_1	τ_2	τ_3	τ_4	β_1	β_2	β_3	β_4
1	738	-4	-1	0	0	-12	27	0	20
2	742	-8	-1	0	0	-16	23	0	16
3	758	-4	-3	0	0	-30	7	-20	0
4	758	-8	-1	0	0	-32	7	-16	0
5	758	-4	0	0	0	-33	7	-20	0
6	738	-4	0	0	0	-13	27	0	20
7	741	-7	0	0	1	-16	23	0	17
8	758	-7	0	0	1	-33	6	-17	0

Exhibit 8. Table of eight optimal extreme point solutions for the LAV estimation problem obtained from exhibit 7. with criterion (3.3) added.

Exhibit 9. Elementary Analysis by Means (i.e. LSQ) of the Nebraska Voting Data. (unit = .1%)

County	'20		'28		'36		'44		'52		'60		eff	fit
D0	15	61	255	88	27	-40	-40	-117	-63	-115	-49	-23	38	387
D1	91	61	8	106	-14	-31	-41	-91	3	-28	-41	-23	-69	280
B1	-29	26	-7	-28	15	14	9	-21	4	8	16	-8	16	365
D2	62	96	-42	-41	2	-20	25	0	-34	-37	-13	2	16	365
D4	26	-49	-47	-20	16	25	42	64	0	5	-41	-22	15	364
B4	-15	47	-33	-46	-13	3	-6	-27	20	31	-3	45	-15	334
D5	-25	-35	-98	-30	-5	30	-27	104	11	44	26	3	-48	301
B5	11	19	-27	-28	-12	27	35	-15	7	30	-17	-29	-42	307
D6	-0	-28	-96	-49	29	5	2	55	16	11	14	42	-36	313
B7	-31	-55	105	-23	-19	-36	-21	-15	-14	10	70	28	53	402
D7	-104	-142	-18	71	-26	23	21	63	49	40	39	-16	73	422
eff	-48	-89	-53	283	130	18	-2	68	-135	-105	-101	32	349	0
fit	301	260	296	632	479	367	347	417	214	244	248	381	0	-349

Exhibit 10. Nebraska voting data from exhibit 2

ROBUST elementary analysis by LAV (unit = .1%)

	Election													
County	'20		'28		'36		'44		'52		'60		eff	fit
D0	56	81	322	144	50	-3	0	-83	-30	-85	-13	0	7	345
D1	124	73	67	154	0	-3	-9	-65	28	-6	-13	-8	-91	247
B1	-25	9	22	-10	0	13	11	-25	-1	0	14	-23	23	361
D2	79	92	0	-10	0	-8	40	10	-26	-32	-1	0	10	348
D4	33	-63	-15	1	4	27	47	64	-1	0	-39	-34	19	357
B4	-7	34	0	-24	-24	6	0	-26	19	27	0	34	-13	325
D5	-9	-40	-57	0	-8	41	-13	112	18	48	36	0	-53	285
B5	13	0	0	-12	-29	24	35	-20	0	20	-20	-46	-33	305
D6	0	-49	-70	-34	10	0	0	48	8	-1	9	24	-25	313
B7	-9	-54	153	14	-15	-18	0	0	0	21	87	32	41	379
D7	-122	-181	-10	68	-63	0	2	38	23	11	16	-52	102	440
eff	-48	-68	-78	268	149	23	0	75	-126	-93	-96	51	338	0
fit	290	270	260	606	487	361	338	413	212	245	242	389	0	-338

Exhibit 11. Analysis of The 132 Residuals Obtained from the LSQ and LAV' Two-way Analyses.

LSQ			LAV'		
4	L	LLLL	2	L	LL
7	-9	751		-9	
	-8		4	-8	35
	-7		5	-7	0
8	-6	3	8	-6	335
9	-5	4	11	-5	472
19	-4	9771580911	14	-4	096
28	-3	043914957	19	-3	40294
47	-2	9487809815062772229	29	-2	5449560603
60	-1	3932169543375	39	-1	0002333558
(7)	-0	7360038	64	-0	7990000833890011060118000
65	0	0335923368524	(20)	0	09001004600002080009
52	1	51516749161254	48	1	34039098946
33	2	669573526	37	2	012344778
29	3	0151090	28	3	3458642
22	4	7255153	21	4	10788
15	5	5	16	5	60
14	6	21143	14	6	784
9	7	21	11	7	93
7	8	9	9	8	17
6	9	16	7	9	2
4	H	HHHH	6	H	HHHHHH

LSQ High	LSQ Low	LAV' High	LAV' Low
256	-104	322	-181
107	-142	154	-122
107	-117	153	
104	-114	144	
		124	
		112	

LSQ		Statistic	LAV'
-27h and +26		Hinges	-13 and +23h
53h		Midspread	36h
-81 and +79		Side Values	-49h and 60
7 and		Number Outside	11 and 14

Exhibit 12 Outside Residuals from row-plus-Column Fits Using LSQ (see Exhibits 9 and 11) and LAD (see Exhibits 10 and 11)

LSQ

	'20.		'28.	.	'36.	.	'44.	.	'52.	.	'60.	.
D0	.	.	255	88	.	.	.	-117	.	-115	.	.
D1	91	.	.	106	.	.	.	-91	.	.	.	.
B1	.	.	.	.	.	.	.	.	.	.	.	.
D2	.	96	.	.	.	.	.	.	.	.	.	.
D4	.	.	.	.	.	.	.	.	.	.	.	.
B4	.	.	.	.	.	.	.	.	.	.	.	.
D5	.	.	-98	.	.	.	.	104	.	.	.	.
B5	.	.	.	.	.	.	.	.	.	.	.	.
D6	.	.	-96	.	.	.	.	.	.	.	.	.
B7	.	.	105	.	.	.	.	.	.	.	.	.
D7	-104	142	.	.	.	.	.	.	.	.	.	.

LAV

	'20.		'28.	.	'36.	.	'44.	.	'52.	.	'60.	.
D0	.	81	322	144	.	.	.	-83	.	-85	.	.
D1	124	73	67	154	.	.	.	-65	.	.	.	.
B1	.	.	.	.	.	.	.	.	.	.	.	.
D2	79	92	.	.	.	.	.	.	.	.	.	.
D4	.	-63	.	.	.	.	.	64	.	.	.	.
B4	.	.	.	.	.	.	.	.	.	.	.	.
D5	.	.	-57	.	.	.	.	112	.	.	.	.
B5	.	.	.	.	.	.	.	.	.	.	.	.
D6	.	.	-70	.	.	.	.	.	.	.	.	.
B7	.	-54	153	.	.	.	.	.	.	.	87	.
D7	-122	-181	.	68	-63	.	.	.	.	.	.	.

Exhibit 13 Elementary Analysis Using Pomedian Procedure On The Nebraska Voting Data (see Tukey [31, exhibit 10 chp. 19]).

	'20		'28		'36		'44		'52		'60		eff	fit
D0	45	91	314	147	59	-14	-9	-73	-36	-95	-19	8	7	356
D1	105	75	53	149	1	-22	-26	-63	14	-2	-27	-8	-83	266
B1	-29	26	21	0	16	9	9	-8	0	-3	14	-8	16	365
D2	62	96	-19	-13	3	-25	-25	14	-38	-48	-13	2	16	365
D4	26	-49	-19	8	17	20	42	78	-3	-6	-41	-22	15	364
B4	-9	51	-1	-14	-8	2	-2	-9	20	24	1	49	-20	329
D5	-24	-34	-69	-1	-3	26	-26	118	8	34	26	4	-49	300
B5	9	13	-1	-2	-13	20	33	-3	1	17	-19	-31	-40	309
D6	0	-28	-67	-20	30	0	2	41	13	0	14	43	-36	313
B7	-30	-54	135	7	-16	-39	-19	0	-16	1	71	30	53	402
D7	-124	-162	-11	24	-45	-2	2	55	22	11	19	-35	93	442
eff	-48	-89	-81	254	129	23	-2	54	-133	-94	-101	32	349	0
Fit	301	260	268	603	478	372	247	403	216	255	248	381	0	-349

REFERENCES

1. Andrews, D.F., "A Robust Method for Multiple Linear Regression" Technometrics 16 (November 1974), 523-31.

2. Armstrong, R.D. and E.L. Frome, "A Comparison of Two Algorithms for Absolute Deviation Curve Fitting" Journal of the American Statistical Association, to appear.

3. Armstrong, R.D. and J.W. Hultz, "A Restricted Discrete Approximation Problem in the L_1 Norm," SIAM Journal on Numerical Analysis, to appear.

4. I. Barrodale, and F.D.K. Roberts, "An Improved Algorithm for Discrete L_1 Linear Approximation," SIAM Journal on Numerical Analysis, 10(1973)839-848.

5. I. Barrodale and A. Young, "Algorithms for Best L_1 and L_∞ Linear Approximation on a Discrete Set," Numerical Mathematics, 8 (1966), 295-306.

6. Beaton, A.E. and Tukey, J.W., "The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data", Technometrics 16 (May 1974), 147-85.

7. A. Charnes, "Optimality and Degeneracy in Linear Programming," Econometrica, 20(1952) 160-170.

8. A. Charnes and W.W. Cooper, "Goal Programming and Constrained Regression -- A Comment", Omega Vol. 3, No. 4 (1975), 403-409.

9. A. Charnes and W.W. Cooper, "Absolute Deviations and Constrained Regressions," Preface to Hildebrand D. translation of De La Vallee Poussin M. Ch. J. "On the Method of Minimum Approximation," Annales de la Societe de Bryxelles, 35, Part II, 1911, 1-16. In ONR Research Memorandum 96, Carnegie Mellon University, Pittsburgh, (1964).

10. A. Charnes, and W.W. Cooper, Management Models and Industrial Applications of Linear Programming, Vol. I and II, John Wiley & Sons, Inc., New YOrk, 1961.

11. A. Charnes, W.W. Cooper and R.J. Nichaus, Studies in Manpower Planning, U.S. Navy Office of Civilian Manpower Management, Washington (1972)

12. A. Charnes, W.W. Cooper and R. Ferguson, "Optimal Estimation of Executive Compensation by Linear Programming," Management Science, 2(1955), 138-151.

13. Gentleman, J.F. and Wilk, M.B., "Detecting Outliers in a Two-Way Table 1. Statistical Behavior of Residuals," Technometrics, 17 (February, 1975), 1-14.

14. Gentleman, J.F. and M.B. Wilk, "Detecting Outliers, II, Supplementing the Direct Analysis of Residuals," Biometrics, 31 (June 1975) 387-410.

15. Glover, F., Karney, D., and Klingman, D., "The Augmented Predecessor Index Method for Locating Stepping Stone Paths and Assigning Dual Prices in Distribution Problems," Transportation Science, Vol. 6 (1972), pp. 171-180.

16. Glover, F., Karney, D., Klingman, D., and Napier, A., "A Computational Study on Start Procedures, Basis Change Criteria, and Solution Algorithms for Transportation Problems." Management Science, Vol. 20, No. 5 (1975), 793-814.

17. Glover, F., D. Klingman and H. Napier, "An Efficient Dual Approach to Network Problems," Opsearch 9, (1972), 1-19.

18. Hadley, G., Linear Programming, Addison-Wesley, Reading, Massachusetts, 1962.

19. Hampel, Frank R., "A General Qualitative Definition of Robustness" Annals of Mathematical Statistics, 42 (No. 6 1971), 1887-96.

20. Harter, H.L., "The Method of Least Squares and Some Alternatives - Part V", International Statistical Review, 43 (No. 3, 1975) 269-78.

21. Hogg, R. V., "Adaptive Robust Procedures; A Partial Review and Some Suggestions For Future Applications and Theory" Journal of the American Statistical Association (December, 1974).

22. Hogg, R. V. and Randles, R. H., "Adaption Distribution - Free Regression Methods and Their Application", Technometrics 17, (November, 1975) 399-407.

23. Huber, P. J., "Robust Statistics: A Review, "The Annals of Mathematical Statistics, 43 (no. 4, 1972) 1041-1067.

24. Lee, S. M., Goal Programming for Decision Analysis, Auerbach, Philadelphia, 1972.

25. Mallows, C. L., "Discussion of Invited Paper" Technometrics, 16 (May 1974) 187-88.

26. McNeill, J. J. and J. W. Tukey "Higher-Order Diagnosis of Two-Way Tables, Illustrated on Two Sets of Demographic Empirical Distributions" Biometrics, (June 1975) 487-510.

27. Prescott, P., "An approximate test for Outliers in Linear Models" Technometrics, (February, 1975), 129-32.

28. P. D. Robers, and A. Ben-Israel, "An Interval Programming Algorithm for Discrete Linear L_1 Approximation Problems," Journal of Approximation Theory, 2(1969), 323-336.

29. Searle, S. R., Linear Models, New York: John Wiley and Sons, 1971.

30. K. Spyropoulos, E. Kiountouzis and A. Young, "Discrete Approximations in the L_1 Norm," The Computer Journal 16(1973) 180-186.

31. Tukey, J. W. Exploratory Data Analysis, (Limited Preliminary Edition), Reading, Mass.: Addison-Wesley, 1970.

32. H. M. Wagner, "Linear Programming Techniques for Regression Analysis," Journal of the American Statistical Association 54 (1959), 206-212.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Center for Cybernetic Studies The University of Texas	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE
(6) Least Absolute Value Estimators for One-Way and Two-Way Tables,

4. DESCRIPTIVE NOTES (Type of report and inclusive dates) (10) (9) Research rept.,

5. AUTHOR(S) (First name, middle initial, last name)
R. D. Armstrong
E. L. Frome

(14) CCS-266

6. REPORT DATE (11) June 1976	7a. TOTAL NO. OF PAGES 36	7b. NO. OF REFS 32

(12) 39p.

8a. CONTRACT OR GRANT NO. (15) N00014-75-C-0569	9a. ORIGINATOR'S REPORT NUMBER(S) Center for Cybernetic Studies Research Report CCS 266
b. PROJECT NO. NR047-021 c. d.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) (16) NR-047-021

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Office of Naval Research (Code 434) Washington, D.C.

13. ABSTRACT

This paper concerns itself with the problem of estimating the parameters in a one-way and two-way classification model by minimizing the sum of the absolute deviations of the regression function from the observed points. The one-way model reduces to obtaining a set of medians from which optimal parameters can be obtained by simple arithmetic manipulations. The two-way model is transformed into a specially structured linear programming problem and two algorithms are presented to solve this problem. The occurrence of alternative optimal solutions in both models is discussed, and numerical examples are presented.

DD FORM 1 NOV 65 1473 (PAGE 1)
S/N 0101-807-6811

406197

Unclassified
Security Classification

A-31408

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Least Absolute Values						
Linear Programming						
Two-Way Tables						
Least Squares						
Robust Regression						

DD FORM 1 NOV 65 1473 (BACK)
S/N 0102-014-6800

Unclassified
Security Classification

A-31409